

Harmonic functions on the Sierpinski triangle

Ilia Smilga

March 4, 2013

Abstract

In this paper, we give a few results on the local behavior of harmonic functions on the Sierpinski triangle - more precisely, of their restriction to a side of the triangle. First we present a general formula that gives the Hölder exponent of such a function in a given point. From this formula, we deduce an explicit algorithm to calculate this exponent in any rational point, and the fact that the derivative of such a function is always equal to 0, ∞ or undefined.

1 Introduction

1.1 Notations and conventions

In all this text, ω stands for the cubic root of -1 with positive imaginary part.

For $z \in \mathbb{C}$, we define $h_z : \mathbb{C} \rightarrow \mathbb{C}$ to be the homothety of center z and ratio $\frac{1}{2}$: $h_z(w) = \frac{w+z}{2}$.

\mathbb{N} stands for the set of all nonnegative integers.

When talking about a positive quantity, we shall say that it is *positively bounded* if it is bounded from above by a finite constant and from below by a positive constant. This is equivalent to saying that its logarithm is bounded.

When we say that some derivative is well-defined, we shall mean that the rate of change has a limit in $\bar{\mathbb{R}}$, i. e. we shall also implicitly admit infinitive derivatives.

We will often need to make estimations up to a positively bounded multiplicative constant. As all norms on a finite-dimensional vector space are equivalent, in such a context, we have no need to distinguish them. So we shall usually simply write “ $\|\bullet\|$ ”, implying that the statement we are making is true for any norm. However, to perform some calculations, we will need to use some specific norms: we shall then use an index to specify which norm we are talking about.

We will also often need to manipulate products of the form $F_{a_1} \dots F_{a_n}$, where F is some kind of operator that can take one of two values F_0 and F_1 (for example h , M , \vec{M} , \tilde{M} and so on). We shall then simply write $F_{a_1 \dots a_n}$, or even F_w (if we have already agreed that $w = a_1 \dots a_n$), as shorthand for such products.

Everything written in positional notation with a radix point is assumed to be in binary (for example $\frac{1}{2} = 0.1$), unless the context clearly shows otherwise (for example $\pi \approx 3.14\dots$).

As everyone knows, dyadic rational numbers have two binary expansions. Being given $s \in [0, 1)$ (resp. $s \in (0, 1]$), we call its *upper* (resp. *lower*) *binary*

expansion the unique binary expansion of s that does not end with an infinite sequence of 1's (resp. of 0's). For the problems that we will study, both expansions will play perfectly symmetrical roles; hence we will often simply say "binary expansion" without specifying, and we shall mean "upper or lower binary expansion".

1.2 Statement of the problem

Let u be the unique continuous function from $[0, 1]$ to \mathcal{R}^3 that satisfies:

$$\begin{cases} u(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & u(1) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ u \circ h_0 = M_0 \circ u \\ u \circ h_1 = M_1 \circ u, \end{cases} \quad (1)$$

where we set

$$M_0 := \frac{1}{5} \begin{pmatrix} 5 & 2 & 2 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad M_1 := \frac{1}{5} \begin{pmatrix} 2 & 0 & 1 \\ 2 & 5 & 2 \\ 1 & 0 & 2 \end{pmatrix}$$

(see also Figure 5). We shall justify later the existence and the unicity of this function.

In [1], Kirillov formulates the question (called "Problem 2", near the end of Section 3.2) :

Compute explicitly the derivative $u'(t)$ whenever it is possible (e.g. at all rational points).

The purpose of our paper is to answer this question. Note however that our notations are slightly different from Kirillov's: what he calls u corresponds in fact to a projection of the function that we have decided to call u . Thus we shall actually deal with a slightly more general question.

1.3 Plan of the paper

We start, in section 2, by explaining the interest of the function u . While mostly following Kirillov's book, we construct harmonic functions on the Sierpinski triangle, and we show that u appears as the restriction of some kind of "universal" harmonic function to a side of the triangle. The rest of the article is logically independent from this section, except for a few notations and definitions. (For a more thorough introduction to harmonic functions on the Sierpinski triangle and on other fractals, see also Strichartz's book [2].)

In section 3, we establish a few general results about that function.

In the following, we study the behaviour of $u(t) - u(s)$ as $t \rightarrow s$. At first (section 4), we completely describe the behaviour of the direction of this vector. This gives the general shape of the curve $u(t)$, and allows us, a few pages further, to link the local properties of the functions that Kirillov studies with those of u .

In section 5, we study the norm of this vector. The key result of this paper is Proposition 9, that links the asymptotic behaviour of this norm with that of an infinite product of matrices indexed on the binary expansion of s .

In section 5.2, we answer the second part of the question: we present an algorithm that allows to calculate the derivative of u (hence also that of $u_a^c{}_b$) in every rational point.

In section 5.3, we answer the question in the general case: we show that the derivative of u , when it exists, can only take the values 0 or ∞ , and that it is almost surely equal to 0.

Finally, in section 5.4, we give some numerical results, which help us make a few estimations. We establish a simple sufficient condition for the derivative to be equal to 0, and we conjecture one for the derivative to be equal to ∞ .

1.4 Acknowledgements

I would like to thank Mr. Yves Benoist, my master's degree advisor, who helped me a lot with this work.

2 Preliminaries : harmonic function on the Sierpinski triangle

Definition 1. Let us first define successive approximations of the Sierpinski triangle: we start with

$$\mathcal{S}_0 := \{0, 1, \omega\},$$

and we define recursively, for all $n \in \mathbb{N}$,

$$\mathcal{S}_{n+1} := h_0(\mathcal{S}_n) \cup h_1(\mathcal{S}_n) \cup h_\omega(\mathcal{S}_n).$$

Let us give a name to the union of all these approximations:

$$\mathcal{S}_\infty := \bigcup_{n=0}^{+\infty} \mathcal{S}_n,$$

and the *Sierpinski triangle* is then defined as the closure of this union:

$$\mathcal{S} := \overline{\mathcal{S}_\infty}.$$

Note that we can naturally see \mathcal{S}_n as a graph: \mathcal{S}_0 is a complete graph on three vertices, and as edges of \mathcal{S}_{n+1} , we take the images of the edges of \mathcal{S}_n by the three homotheties. Then it is easy to see that every vertex except 0, 1 and ω has exactly four neighbours. (See Figure 1.) In the following, E is a real vector space, and we will write “ $s \sim t$ ” for “ s is a neighbour of t ”.

Definition 2. Let $n \in \mathbb{N}$. A *harmonic function* on \mathcal{S}_n is a function $f : \mathcal{S}_n \rightarrow E$ such that:

$$\forall s \in \mathcal{S}_n \setminus \{0, 1, \omega\}, \quad f(s) = \frac{1}{4} \sum_{t \sim s} f(t).$$

We impose no condition on the values $f(0)$, $f(1)$ and $f(\omega)$: we consider them to be boundary conditions.

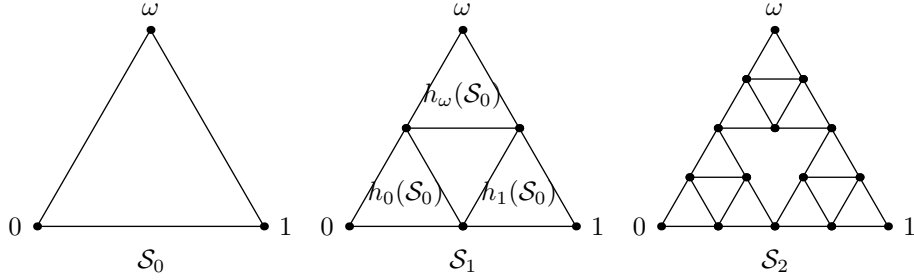


Figure 1: Discrete approximations of Sierpinski triangle.

Definition 3. A *harmonic function* on \mathcal{S} is a continuous function $f : \mathcal{S} \rightarrow E$ whose restriction to every \mathcal{S}_n is harmonic.

Proposition-definition 1. Let $a, b, c \in E$. Then there exists a unique function f that is harmonic on \mathcal{S} and that satisfies the boundary conditions:

$$f(0) = a, \quad f(1) = b, \quad f(\omega) = c.$$

(This means in particular that the set of all real-valued harmonic functions on \mathcal{S} is a three-dimensionnal vector space.) We will call this function $f_a^c_b$.

Proof. To check this, we will explicitly construct, by induction, the family of functions that verify these conditions on successive approximations of \mathcal{S} .

First we define the (obviously unique) harmonic function on $\mathcal{S}_0 = \{0, 1, \omega\}$ that satisfies the boundary conditions:

$$f_0 : \quad 0 \mapsto a, \quad 1 \mapsto b, \quad \omega \mapsto c.$$

Now suppose we have a harmonic function $f_n : \mathcal{S}_n \rightarrow E$ satisfying the boundary conditions. We will now define a function f_{n+1} that extends f_n to \mathcal{S}_{n+1} . Let s, t, u be any three adjacent vertices of \mathcal{S}_n that form a triangle having the same orientation as \mathcal{S}_0 . Then \mathcal{S}_{n+1} still contains these vertices, but also the vertices $\frac{s+t}{2}, \frac{t+u}{2}, \frac{s+u}{2}$ (see Figure 2). Conversely, it is easy to see that all vertices of \mathcal{S}_{n+1} can be accounted for in this way. We then set

$$\begin{aligned} f_{n+1}\left(\frac{s+t}{2}\right) &= \frac{2f_n(s) + 2f_n(t) + f_n(u)}{5} \\ f_{n+1}\left(\frac{s+u}{2}\right) &= \frac{2f_n(s) + f_n(t) + 2f_n(u)}{5} \\ f_{n+1}\left(\frac{t+u}{2}\right) &= \frac{f_n(s) + 2f_n(t) + 2f_n(u)}{5} \end{aligned}$$

Being an extension of f_n , this function still satisfies the boundary conditions. Checking that it is harmonic on \mathcal{S}_{n+1} is a straightforward calculation; it is enough to do it in s and in $\frac{s+t}{2}$, since any vertex is similar to one of those two.

- The neighbours of s in \mathcal{S}_{n+1} are $\frac{s+t}{2}, \frac{s+u}{2}, \frac{s+t'}{2}$ and $\frac{s+u'}{2}$, where t' and u'

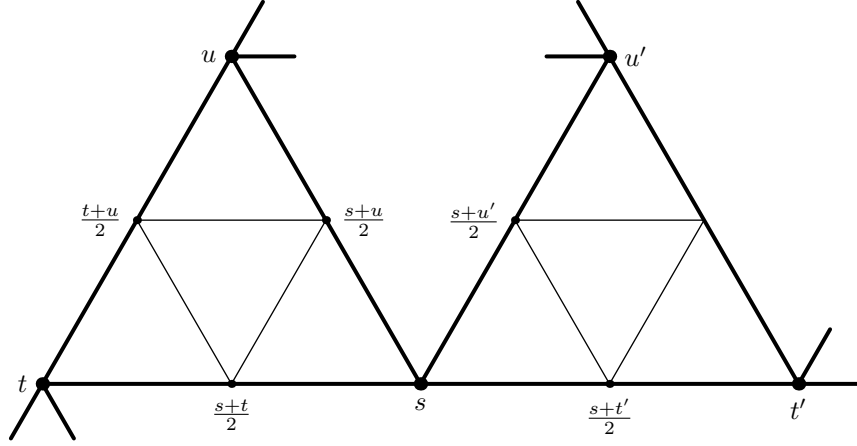


Figure 2: Subdivision of \mathcal{S}_n .

are the two other neighbours of s in \mathcal{S}_n (see Figure 2). We have:

$$\begin{aligned}
& f_{n+1}\left(\frac{s+t}{2}\right) + f_{n+1}\left(\frac{s+u}{2}\right) + f_{n+1}\left(\frac{s+t'}{2}\right) + f_{n+1}\left(\frac{s+u'}{2}\right) = \\
&= \frac{2f_n(s) + 2f_n(t) + f_n(u)}{5} + \frac{2f_n(s) + f_n(t) + 2f_n(u)}{5} + \\
&\quad + \frac{2f_n(s) + 2f_n(t') + f_n(u')}{5} + \frac{2f_n(s) + f_n(t') + 2f_n(u')}{5} \\
&= \frac{8f_n(s) + 3(f_n(t) + f_n(u) + f_n(t') + f_n(u'))}{5} \\
&= \frac{8 + 12}{5} f_n(s) \\
&= 4f_{n+1}(s)
\end{aligned}$$

- The neighbours of $\frac{s+t}{2}$ are s , $\frac{s+u}{2}$, $\frac{t+u}{2}$ and t . We have:

$$\begin{aligned}
& f_{n+1}(s) + f_{n+1}\left(\frac{s+u}{2}\right) + f_{n+1}\left(\frac{t+u}{2}\right) + f_{n+1}(t) = \\
&= f_n(s) + \frac{2f_n(s) + f_n(t) + 2f_n(u)}{5} + \frac{f_n(s) + 2f_n(t) + 2f_n(u)}{5} + f_n(t) \\
&= \frac{8f_n(s) + 8f_n(t) + 4f_n(u)}{5} \\
&= 4f_{n+1}\left(\frac{s+t}{2}\right)
\end{aligned}$$

Now let us show that these functions are unique. Indeed, consider two harmonic functions on \mathcal{S}_n that satisfy the same boundary conditions. This means that their difference vanishes at 0, 1 and ω , and is harmonic as well. Then the maximum principle guarantees that it is identically equal to zero.

This allows us to define $f_a^c_b$ on \mathcal{S}_∞ , and the above calculation gives us an explicit recurrence relation to compute it in any point. Let us rewrite this relation in a more convenient way:

$$\begin{cases} f_a^c_b \circ h_0 = f_a^{\frac{2a+b+2c}{5}}_{\frac{2a+2b+c}{5}} \\ f_a^c_b \circ h_1 = f_{\frac{2a+2b+c}{5}}^{\frac{a+2b+2c}{5}}_b \\ f_a^c_b \circ h_\omega = f_{\frac{2a+b+2c}{5}}^c_{\frac{a+2b+2c}{5}}. \end{cases} \quad (2)$$

To extend this function continuously to the whole Sierpinski triangle, we need to check that it is uniformly continuous. Let

$$D(a, b, c) := \max(\|a - b\|, \|b - c\|, \|a - c\|).$$

By the maximum principle, we have, for all values of a, b, c and for all $s, t \in \mathcal{S}_\infty$,

$$\|f_a^c_b(s) - f_a^c_b(t)\| \leq D(a, b, c).$$

On the other hand, it is easy to check that we have

$$D(a, \frac{2a+b+2c}{5}, \frac{2a+2b+c}{5}) \leq \frac{3}{5} D(a, b, c).$$

Using (2), by induction, it follows that if T is an “elementary triangle” of level n , we have

$$\forall s, t \in T, \quad \|f_a^c_b(s) - f_a^c_b(t)\| \leq \left(\frac{3}{5}\right)^n C,$$

where C depends only on a, b and c . Now consider two points of \mathcal{S}_∞ whose distance is at most $(\frac{1}{2})^n$; then it is easy to see that the respective elementary triangles of level $n - 1$ that contain them are either coincident or adjacent. It follows that

$$\forall s, t \in \mathcal{S}_\infty, \quad |s - t| \leq \left(\frac{1}{2}\right)^n \implies \|f_a^c_b(s) - f_a^c_b(t)\| \leq \left(\frac{3}{5}\right)^n C',$$

where C' is still some positive real constant that depends only on a, b and c — which leads to the conclusion. This shows that $f_a^c_b$ is indeed well-defined on \mathcal{S} . \square

Very informally, $f_a^c_b$ describes the shape that a Sierpinski triangle made of rubber would assume if it were stretched between three nails fixed in a, b and c (see Figure 3).

Note that by continuity, the relations (2) are still valid in \mathcal{S} . They are very important: they supply an intrinsic definition of the functions $f_a^c_b$.

From now on, following Kirillov, we restrict ourselves to a side of the Sierpinski triangle: we let $u_a^c_b = f_a^c_b|_{[0,1]}$. We may then drop the third line from (2). Moreover, Kirillov is mainly interested in real-valued functions, in particular four of them: $\phi = u_0^0_1$, $\psi = u_0^1_1$, $\chi = u_0^{-1}_1$ and $\xi = u_0^2_1$ (in his notations, u actually stands for some one function among these four). However, they have a drawback: it is difficult to study one of this functions independently from others, because (2) forces us to change the values of a, b and c .

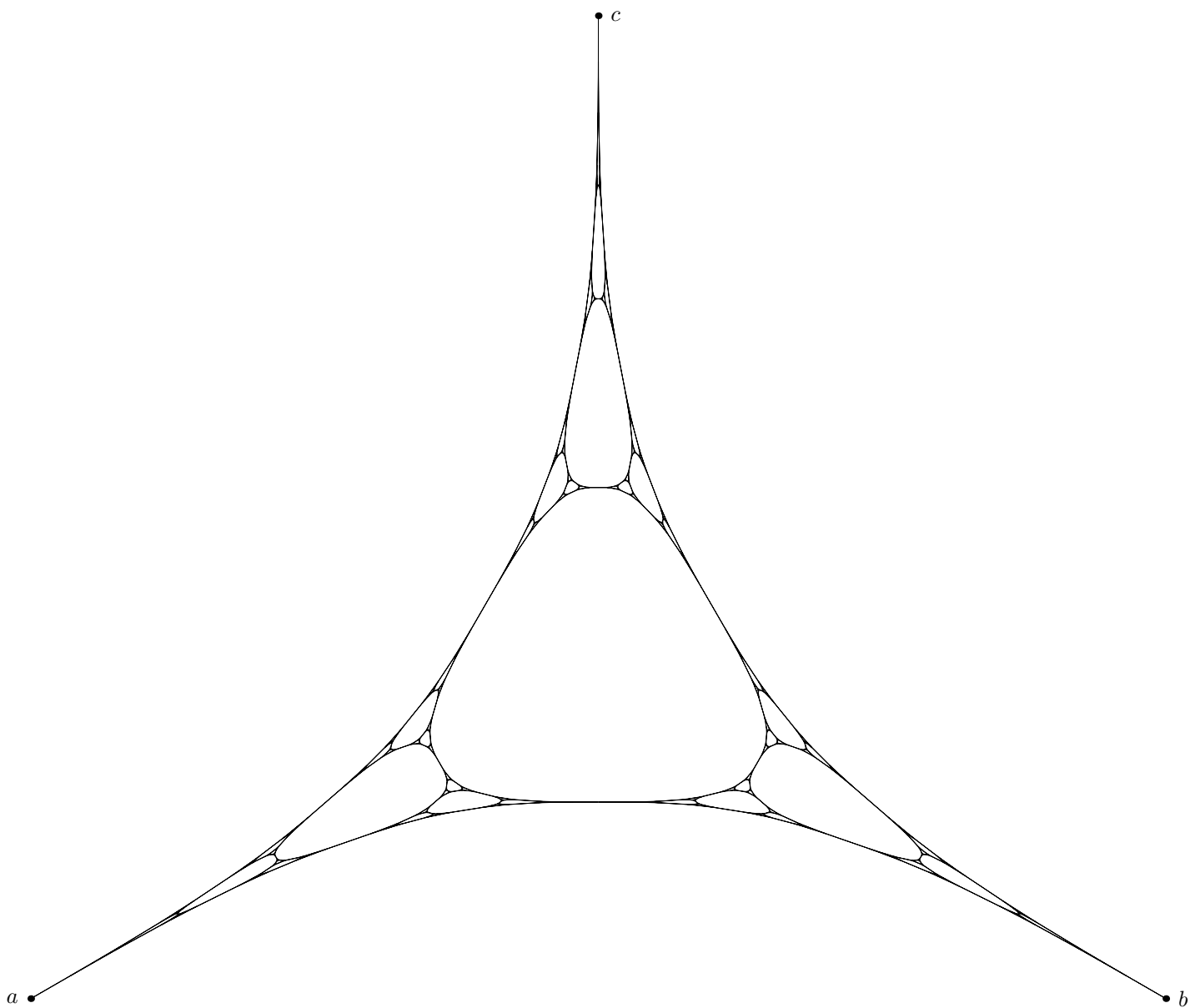


Figure 3: Harmonic image of the Sierpinski triangle, $f_a^c_b(\mathcal{S})$.

This leads us to introduce the function $u_{e_0^{e_\omega} e_1}$, where (e_0, e_1, e_ω) is the canonical basis of \mathbb{R}^3 . Its main interest is that it contains all the information about all the functions $u_a^c b$. Indeed, for all a, b and c , we have $u_a^c b = \Phi \circ u_{e_0^{e_\omega} e_1}$, where $\Phi : \mathbb{R}^3 \rightarrow E$ is the linear map that send the canonical basis to (a, b, c) . (If a, b and c are real, it is simply the linear form with matrix $\begin{pmatrix} a & b & c \end{pmatrix}$.) The relations (2) then give us a functional equation on $u_{e_0^{e_\omega} e_1}$ — which is none other than (1). Thus we see that the function u that we have defined in the beginning is just a shorter notation (that we shall adopt from now on) for $u_{e_0^{e_\omega} e_1}$. Its existence and unicity follow from the proof of Proposition-Definition 1.

3 General properties of u

Note that all that we shall say in this section can very easily be generalized to all of \mathcal{S} , that is to $f_{e_0^{e_\omega} e_1}$. To do this, it is enough to introduce the matrix M_ω , analogous to M_0 and M_1 , and to see that a generic point of \mathcal{S} can be described by an infinite sequence of symbols 0, 1 or ω , by analogy with the binary expansion.

We will need to diagonalise the M_i . Any of these matrices has eigenvalues 1, $\frac{3}{5}$ and $\frac{1}{5}$, with respective eigenvectors e_i , \vec{v}_i and \vec{w}_i , where:

$$e_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{v}_0 = \begin{pmatrix} -1 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \vec{w}_0 = \begin{pmatrix} 0 \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix},$$

$$e_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_1 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{pmatrix}, \quad \vec{w}_1 = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{pmatrix}.$$

Note that

$$f_1^0_0 + f_0^0_1 + f_0^1_0 = f_1^1_1 \equiv 1,$$

which shows that our picture is in fact only two-dimensionnal: all the values of $f_{e_0^{e_\omega} e_1}$, hence also all the values of u , actually lie in the affine plane

$$\mathcal{H} := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1\}.$$

Of course, it is stable by the maps M_i , and the latter induce some affine maps on the former (see Figure 4). In fact, this is why we have chosen to write e_i instead of \vec{e}_i : the e_i are elements of the affine plane \mathcal{H} , namely — under this interpretation — fixed points of the affine maps M_i .

Let us now try to write down an explicit formula for $u(s)$. We first need to express s in terms of 0, of 1 and of the h_i , which may be done by writing:

$$s = \lim_{n \rightarrow \infty} h_{a_1 \dots a_n}(0) = \lim_{n \rightarrow \infty} h_{a_1 \dots a_n}(1),$$

where $s = 0, a_1 a_2 \dots$ is a (lower or upper) binary expansion of s . Using the above relations, we get the following:

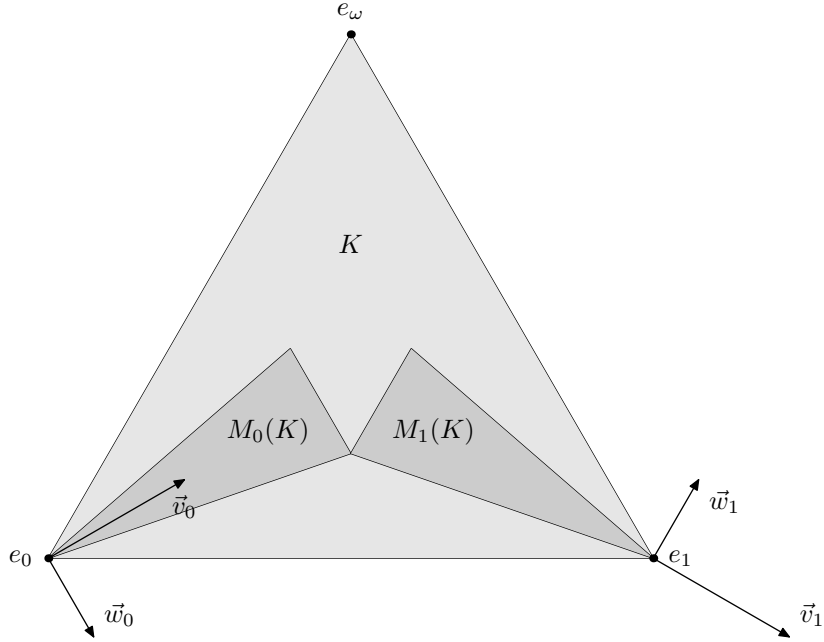


Figure 4: Eigenvectors of the operators M_i ; the plane of the sheet of paper corresponds to the affine plane \mathcal{H} . Thus the origin of \mathbb{R}^3 lies beneath the centre of the figure. The triangle K is defined a little bit further; it is added here to better illustrate the action of the M_i on \mathcal{H} .

Proposition 2. *Let $K := \mathcal{H} \cap \mathbb{R}_{\geq 0}^3$ be the (full) triangle with the canonical coordinate vectors as vertices. Then for all $s \in [0, 1]$ and for any initial value $u_0 \in K$, we have*

$$u(s) = \lim_{n \rightarrow \infty} M_{a_1 \dots a_n}(u_0), \quad (3)$$

where $s = 0.a_1 a_2 \dots$ is its (lower or upper) binary expansion.

(See Figure 5.)

Proof. Let us first rewrite this formula in such a way as to make it clear that the result does not depend on the initial value. It is easy to check that K is stable by both M_0 and M_1 : thus the sequence $M_{a_1 \dots a_n}(K)$ is decreasing. In the proof of Proposition-Definition 1, we have already seen that the diameters of these sets tend to 0. It is clear that they are closed. Thus their intersection contains a single point:

$$\forall u_0 \in K, \quad \left\{ \lim_{n \rightarrow \infty} M_{a_1 \dots a_n}(u_0) \right\} = \bigcap_{n=0}^{\infty} M_{a_1 \dots a_n}(K).$$

On the other hand, a trivial induction shows that for any finite word w on the alphabet $\{0, 1\}$, $h_w(0) = 0.w$, where $0.w$ is seen as a binary fraction, and that $u \circ h_w = M_w \circ u$. Hence, by the relations (1) and by continuity of u ,

$$u \left(\sum_{i=1}^{\infty} a_i 2^{-i} \right) = \lim_{n \rightarrow \infty} M_{a_1 \dots a_n} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

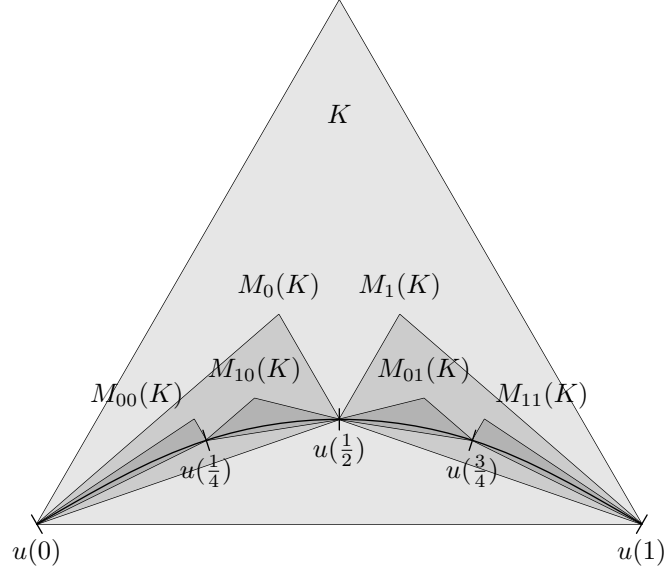


Figure 5: Illustration of Proposition 2. The curve that runs through the picture is the parametric curve $u(t)$.

and hence the formula. \square

Corollary 3. u is injective.

Proof. Let $(a_i)_{i \geq 1}$ and $(a'_i)_{i \geq 1}$ be two sequences with values in $\{0, 1\}$. Let k be the first index where these sequences diverge, i. e. suppose $a_1 = a'_1, \dots, a_{k-1} = a'_{k-1}$, $a_k \neq a'_k$. Without loss of generality, we may actually suppose that $a_k = 0$ and $a'_k = 1$. Choose some $u_0 \in K$. Then the following statements are equivalent (remember that the M_i are bijective):

$$\begin{aligned} \lim_{n \rightarrow \infty} M_{a_1 \dots a_n}(u_0) &= \lim_{n \rightarrow \infty} M_{a'_1 \dots a'_n}(u_0) \\ \bigcap_{n=0}^{\infty} M_{a_1 \dots a_n}(K) &= \bigcap_{n=0}^{\infty} M_{a'_1 \dots a'_n}(K) \\ \bigcap_{n=k}^{\infty} M_{0a_{k+1} \dots a_n}(K) &= \bigcap_{n=k}^{\infty} M_{0a'_{k+1} \dots a'_n}(K) \end{aligned}$$

On the last line, each side is a subset of the respective $M_i(K)$. But a quick calculation (or a quick look at Figure 5) shows that the two images of the triangle intersect at a single point:

$$M_0(K) \cap M_1(K) = \left\{ \frac{1}{5} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \right\} = \{M_0(e_1)\} = \{M_1(e_0)\}.$$

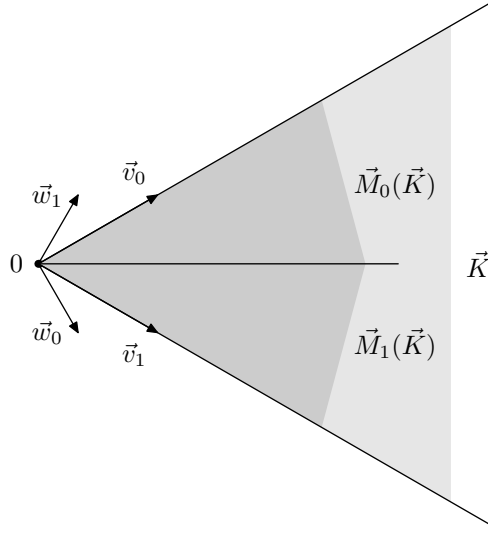


Figure 6: Main objects of the vector plane $\vec{\mathcal{H}}$.

Thus the previous equality holds if and only if

$$\begin{cases} \bigcap_{n=k}^{\infty} M_{a_{k+1} \dots a_n}(K) = \{e_1\} \\ \bigcap_{n=k}^{\infty} M_{a'_{k+1} \dots a'_n}(K) = \{e_0\}, \end{cases}$$

which is true if and only if for all $i > k$, $a_i = 1$ and $a'_i = 0$. In other words, both sequences yield the same value if and only if we have $\sum_{i=1}^{\infty} a_i 2^{-i} = \sum_{i=1}^{\infty} a'_i 2^{-i}$. The “only if” part proves our claim, and the “if” part offers additional support for the equivalence of different binary expansions. \square

4 Local behavior of u : existence and direction of tangent

We will now study the behaviour of the vector $u(t) - u(s)$ as $t \rightarrow s$. In this section, we will thoroughly describe the behaviour of the norm of this vector: quantitatively through the formula (4), and qualitatively in Corollary 6.

The difference $u(t) - u(s)$ lives in $\vec{\mathcal{H}} := \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$, the vector plane parallel to the affine plane \mathcal{H} . This vector plane is of course stable by M_0 and M_1 ; it is thus natural to introduce the restrictions of these operators to $\vec{\mathcal{H}}$, that we shall call \vec{M}_0 and \vec{M}_1 . Each \vec{M}_i has eigenvalues $\frac{3}{5}$ and $\frac{1}{5}$, with respective eigenvectors \vec{v}_i and \vec{w}_i . Note that in both cases, \vec{v}_i and \vec{w}_i actually form an orthogonal basis of $\vec{\mathcal{H}}$.

Let us now introduce the set

$$\vec{K} := \{a\vec{v}_0 + b\vec{v}_1 \mid a, b \geq 0 \text{ and } (a, b) \neq (0, 0)\}$$

(see Figure 6.) This notation should not lead the reader to think that this set is somehow derived from K : it is not, at least not directly. But it does play an analogous role, and the following lemma will show precisely in what sense:

Lemma 4. *The set \vec{K} is stable by the \vec{M}_i . For all $s, t \in [0, 1]$ such that $s < t$, $u(t) - u(s) \in \vec{K}$.*

Proof. The first part can be checked by a straightforward calculation (see also Figure 6).

As for the second part, let us first prove the result for dyadic rational values of s and t (such that $0 \leq s < t \leq 1$), by induction on their dyadic valuations. Let us write $I_n := \mathcal{S}_n \cap [0, 1] = \{\frac{k}{2^n} \mid 0 \leq k \leq 2^n\}$, so that we have $I_0 = \{0, 1\}$ and $I_{n+1} = h_0(I_n) \cup h_1(I_n)$. If $s, t \in I_0$, then $s = 0, t = 1$ and the conclusion can be checked by a straightforward calculation. Now suppose that the conclusion is true for all $s, t \in I_n$; let $s < t$ be two points in I_{n+1} . We need to distinguish three cases: $s < t \leq \frac{1}{2}$, $s < \frac{1}{2} < t$ and $\frac{1}{2} \leq s < t$. We shall only treat the middle case, which is the hardest one; the other two are analogous. We may write $s = h_0(S)$ and $t = h_1(T)$, so that we have:

$$\begin{aligned} u(t) - u(s) &= u(t) - u(\tfrac{1}{2}) + u(\tfrac{1}{2}) - u(s) \\ &= u(h_1(T)) - u(h_1(0)) + u(h_0(1)) - u(h_0(S)) \\ &= \vec{M}_1(u(T) - u(0)) + \vec{M}_0(u(1) - u(S)). \end{aligned}$$

Applying, in order, the induction hypothesis, the fact that \vec{K} is stable by the \vec{M}_i , and the fact that \vec{K} is stable under linear combination with positive coefficients (which follows immediately from the formula that defines it), we get the result.

Now let $0 \leq s < t \leq 1$ be general values. Since u is continuous and since dyadic rationals are dense, $u(t) - u(s)$ lies in the closure of \vec{K} , which is equal to $\vec{K} \cup \{0\}$. But since, by Corollary 3, u is injective, the difference cannot be equal to zero. \square

We now need to introduce the projective equivalents of the relevant objects. Let $\pi : \vec{\mathcal{H}} \setminus \{0\} \rightarrow \mathbb{P}^1(\mathbb{R})$ be the canonical projection; we let $\tilde{K} := \pi(\vec{K})$ and

$$\begin{aligned} \tilde{M}_i : \quad \mathbb{P}^1(\mathbb{R}) &\rightarrow \mathbb{P}^1(\mathbb{R}) \\ \pi(\vec{x}) &\mapsto \pi(M_i(\vec{x})) \end{aligned}$$

(see Figure 7). Clearly, \tilde{K} is stable by the \tilde{M}_i .

Proposition 5. *Let $s \in [0, 1]$; let $s = 0, a_1 a_2 \dots$ be its upper binary expansion. Then for all $\tilde{u}_0 \in \tilde{K}$, we have*

$$\lim_{\substack{t \rightarrow s \\ t > s}} \pi(u(t) - u(s)) = \lim_{n \rightarrow \infty} \tilde{M}_{a_1 \dots a_n}(\tilde{u}_0), \quad (4)$$

that is, both limits always exist and are always equal.

As for lower expansions, the formula becomes:

$$\lim_{\substack{t \rightarrow s \\ t < s}} \pi(u(s) - u(t)) = \lim_{n \rightarrow \infty} \tilde{M}_{a_1 \dots a_n}(\tilde{u}_0).$$

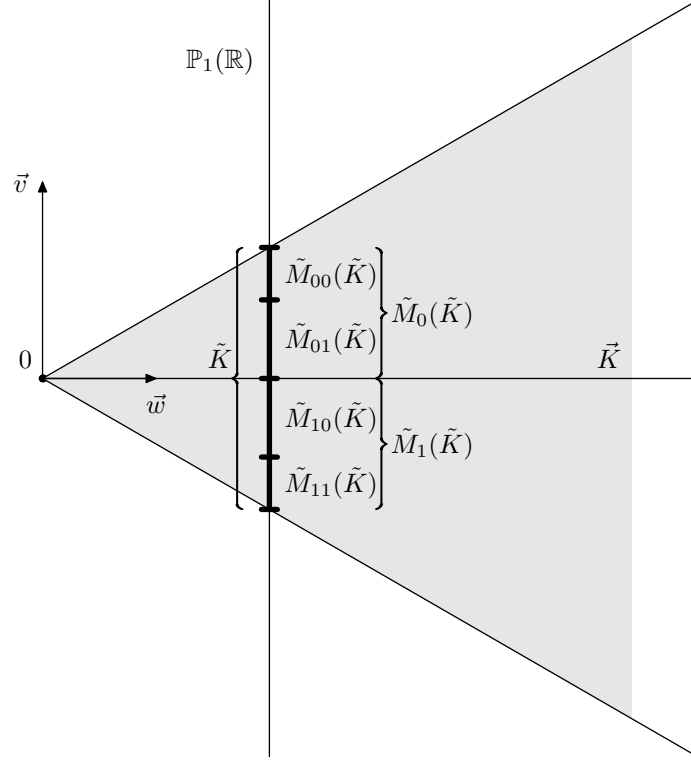


Figure 7: Illustration of Proposition 5

Proof. The proof is quite similar to the proof of Proposition 2. Just as before, the sets $\tilde{M}_{a_1 \dots a_n}(\tilde{K})$ are all closed and form a decreasing sequence. To ensure that their intersection contains a single point, it is enough to check that their diameters tend to 0; to check this, it is enough to show that the \tilde{M}_i are contractive. In the usual metric (where the distance between two points is the angle between the corresponding lines), this is not the case: the Lipschitz constant is equal to 1. But we shall use another metric — the one defined by

$$d(\pi(\vec{u}_1), \pi(\vec{u}_2)) = \left| \frac{a_1}{b_1} - \frac{a_2}{b_2} \right|, \text{ where we let } \vec{u}_i = a_i \vec{v} + b_i \vec{w} \text{ with } \vec{v} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} \text{ and}$$

$$\vec{w} = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix}. \text{ Informally, what we do is to project on a line rather than on a}$$

circle: see Figure 7. Since there exists a neighbourhood of the point at infinity which does not intersect \tilde{K} , this metric is indeed defined on \tilde{K} and induces the right topology on it. Then a few lines of calculation show that in this metric, both \tilde{M}_0 and \tilde{M}_1 have Lipschitz constant $\frac{3}{4}$. Hence we can write

$$\forall \tilde{u}_0 \in \tilde{K}, \quad \left\{ \lim_{n \rightarrow \infty} \tilde{M}_{a_1 \dots a_n}(\tilde{u}_0) \right\} = \bigcap_{n=0}^{\infty} \tilde{M}_{a_1 \dots a_n}(\tilde{K}).$$

Now let $n \in \mathbb{N}$. Then for any t sufficiently close to, but greater than, s , the first n bits of the binary expansions of t and s coincide (this is true because

we use the upper binary expansion of s), so that we have $s = h_{a_1 \dots a_n}(S)$ and $t = h_{a_1 \dots a_n}(T)$ for some $S, T \in [0, 1]$. By Lemma 4, we have

$$\pi(u(t) - u(s)) = \tilde{M}_{a_1 \dots a_n}(\pi(u(T) - u(S))) \in \tilde{M}_{a_1 \dots a_n}(\tilde{K}),$$

hence the result.

As for the symmetric case, the proof is completely analogous. \square

Corollary 6. *The quantity*

$$\vec{D}u(s) := \lim_{\substack{t \rightarrow s \\ t \in [0, 1]}} \pi(u(t) - u(s)) \operatorname{sgn}(t - s)$$

is a well-defined, injective, continuous function of s .

In geometric terms, $\vec{D}u(s)$ is the direction of the tangent to the curve which is the *image* of u . This result may be interpreted as saying that this curve is C^1 and convex — which is after all not so surprising if we think of it as the shape of the side of an elastic Sierpinski triangle stretched on three nails (see Figure 3). Note that the *parametrization* of the curve is much less regular than that — as we shall see, it is not even always differentiable.

Proof. Being given two sequences $(a_i)_{i \geq 1}$ and $(a'_i)_{i \geq 1}$ with values in $\{0, 1\}$, we claim that

$$\bigcap_{n=0}^{\infty} \tilde{M}_{a_1 \dots a_n}(\tilde{K}) = \bigcap_{n=0}^{\infty} \tilde{M}_{a'_1 \dots a'_n}(\tilde{K})$$

if and only if both sequences are binary expansions of the same number. The proof simply follows, *mutatis mutandis*, the proof of Corollary 3; we need the fact that the \tilde{M}_i are injective and that $\tilde{M}_0(\tilde{K})$ and $\tilde{M}_1(\tilde{K})$ intersect at a single point, which are easily seen to be true. Combined with the previous Proposition, the “if” part shows that the function is well-defined and the “only if” part shows that it is injective.

As for the continuity, let $\varepsilon > 0$. Since \tilde{M}_0 and \tilde{M}_1 are contractive (in a suitable metric), we can find a number n such that the diameter of every possible $\tilde{M}_{a_1 \dots a_n}(\tilde{K})$ is less than $\frac{\varepsilon}{2}$. Now let $s, s' \in [0, 1]$ such that $|s - s'| \leq 2^{-n}$. Let $s = 0.a_1 \dots a_n \dots$ and $s' = 0.a'_1 \dots a'_n \dots$ be their binary expansions; let $w = a_1 \dots a_n$, so that $\vec{D}u(s) \in \tilde{M}_w(\tilde{K})$, and similarly define w' . Then two cases are possible: either $w = w'$, or w and w' are two consecutive words. However, knowing that $\tilde{M}_0(\pi(\vec{w}_0)) = \pi(\vec{w}_0)$, $\tilde{M}_1(\pi(\vec{w}_1)) = \pi(\vec{w}_1)$ and $\tilde{M}_1(\pi(\vec{w}_0)) = \tilde{M}_0(\pi(\vec{w}_1))$, a trivial induction shows that in the latter case, $\tilde{M}_w(\tilde{K})$ and $\tilde{M}_{w'}(\tilde{K})$ have a common point. Hence the distance between $\vec{D}u(s)$ and $\vec{D}u(s')$ is less than ε , which finishes the proof. \square

5 Local behavior of u and $u_a^c{}_b$: rate of change

5.1 Hölder exponent formula

In the previous section, we have seen that as t tends to s , the direction of $u(t) - u(s)$ behaves rather well. We shall now see that its norm behaves much more erratically. A naive approach would be to simply try to calculate the

derivative $u'(s)$; however, we shall see in Proposition 16 that whenever this derivative exists, it is equal to either 0 or ∞ , so it is not the most relevant parameter here.

We introduce the following parameters, that will provide a better description of the local behavior of u :

Definition 4. Let $s \in [0, 1]$.

- The *upper Hölder exponent* of u in s is

$$\alpha_{\sup}(u, s) := \limsup_{\substack{t \rightarrow s \\ t \in [0, 1]}} \frac{\ln \|u(t) - u(s)\|}{\ln |t - s|}.$$

- The *lower Hölder exponent* of u in s is

$$\alpha_{\inf}(u, s) := \liminf_{\substack{t \rightarrow s \\ t \in [0, 1]}} \frac{\ln \|u(t) - u(s)\|}{\ln |t - s|}.$$

- When both are equal, we shall simply call them the *Hölder exponent* of u in s , and we shall write

$$\alpha(u, s) := \lim_{\substack{t \rightarrow s \\ t \in [0, 1]}} \frac{\ln \|u(t) - u(s)\|}{\ln |t - s|}.$$

Once we know the Hölder exponent of u , we can usually calculate its derivative, and then the derivative of $u_a^c b$. Indeed:

Lemma 7. Let $s \in [0, 1]$. Then we have :

- $\alpha_{\inf}(u, s) > 1 \implies u'(s) = 0 \implies \alpha_{\inf}(u, s) \geq 1$;
- $\alpha_{\sup}(u, s) < 1 \implies u'(s) = \infty \implies \alpha_{\sup}(u, s) \leq 1$,

where the latter equality is shorthand for “ $\lim_{t \rightarrow s} \frac{\|u(t) - u(s)\|}{|t - s|} = \infty$ ”.

Proof. This is an easy exercise in real analysis. \square

Lemma 8. Let $\Phi = (a \ b \ c)$ be a linear form, and let $s \in [0, 1]$ such that $\vec{D}u(s) \notin \ker \Phi$. Then we have:

- $(u_a^c b)'(s) = 0$ iff $u'(s) = 0$;
- $(u_a^c b)'(s) = \infty$ iff $u'(s) = \infty$;
- $(u_a^c b)'(s)$ is undefined iff $u'(s)$ is undefined.

Proof. This is basically an application of the formula for the derivative of a composite function. \square

Note that by Corollary 6, for a given value of Φ , this lemma applies for all values of s except at most one. In particular, it is easy to check that for the functions ϕ , ψ , χ and ξ , introduced in [1], this lemma holds with only two exceptions: $\chi'(0)$ and $\xi'(1)$ (and in [1], it is shown that both are equal to 0.)

Here is the formula that allows one to calculate the Hölder exponent:

Proposition 9. *Let $s \in [0, 1]$, $s = 0.a_1a_2\dots$ its binary expansion. Then we have:*

$$\begin{aligned}\alpha_{\inf}(u, s) &= \liminf_{n \rightarrow \infty} \frac{\ln \|\vec{M}_{a_1 \dots a_n}\|}{n \ln \left(\frac{1}{2}\right)}, \\ \alpha_{\sup}(u, s) &= \limsup_{n \rightarrow \infty} \frac{\ln \|\vec{M}_{a_1 \dots a_n}\|}{n \ln \left(\frac{1}{2}\right)}\end{aligned}\tag{5}$$

(compare this with (3) and (4)). Before proving it, we shall need the following lemma:

Lemma 10. *Let f be a linear automorphism of \mathbb{R}^2 , let \mathcal{C} be a salient closed cone that is stable by f , and let $\mathcal{C}' \subset \text{int } \mathcal{C}$ such that $\mathcal{C}' \cup \{0\}$ is also a closed cone. Then for all $\vec{x} \in \mathcal{C}'$, the ratio*

$$\frac{\|f(\vec{x})\|}{\|f\| \|\vec{x}\|}$$

is positively bounded by constants that depend only on \mathcal{C} and \mathcal{C}' .

Proof. Let us choose a basis in which \mathcal{C} becomes the upper right quadrant. We may now write:

- $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}$;
- $f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where the coefficients a, b, c and d are all nonnegative;
- $\mathcal{C}' = \left\{ (x, y) \in \mathcal{C} \setminus \{0\} \mid m \leq \frac{x}{y} \leq M \right\}$, where m and M are some finite positive bounds.

Let $\vec{x} = (x, y) \in \mathcal{C}'$. Let us use the supremum norm: we have $\|\vec{x}\|_{\infty} = \max(x, y)$, $\|f\|_{\infty} = \max(a, b, c, d)$ and $\|f(\vec{x})\|_{\infty} = \max(ax + by, cx + dy)$. But we have, on the one hand

$$\max(ax + by, cx + dy) \leq 2 \max(a, b, c, d) \max(x, y),$$

and on the other hand

$$\begin{aligned}\max(ax + by, cx + dy) &\geq \max(a, b, c, d) \min(x, y) \\ &\geq \mu \max(a, b, c, d) \max(x, y),\end{aligned}$$

where $\mu = \min(m, M^{-1})$ depends only on \mathcal{C} and \mathcal{C}' as announced. \square

Let $\vec{K}' = \{a\vec{w}_0 + b\vec{w}_1 \mid a, b \geq 0\}$ (on Figure 6, it is the cone that lies between \vec{w}_0 and \vec{w}_1 , not represented to avoid overloading the picture). Then it is easy to check that it is stable by \vec{M}_0 and \vec{M}_1 , hence also by all of their products. By applying the lemma on $\mathcal{C} = \vec{K}'$ and $\mathcal{C}' = \vec{K}$ (watch out for the slight inconsistency in notations), we get the following estimation:

Corollary 11. *For any finite word w on the alphabet $\{0, 1\}$ and for all $\vec{u}_0 \in \vec{K}$, the ratio*

$$\frac{\|\vec{M}_w(\vec{u}_0)\|}{\|\vec{M}_w\| \|\vec{u}_0\|}$$

is positively bounded.

Proof of Proposition 9. Now let $s \in [0, 1]$, and let $s' \in [0, 1]$ be a point close to, but not equal to, s . Let $n = \left\lfloor \frac{\ln |s' - s|}{\ln(\frac{1}{2})} \right\rfloor$, so that $\frac{1}{2} \leq 2^n |s' - s| \leq 1$. Let $s = 0.a_1 \dots a_n \dots$ and $s' = 0.a'_1 \dots a'_n \dots$ be the respective binary expansions. Let us write $w = a_1 \dots a_n$ and $w' = a'_1 \dots a'_n$; we then have $s = h_w(S)$ and $s' = h_{w'}(S')$, for some $S, S' \in [0, 1]$, hence:

$$\begin{aligned} \frac{\ln \|u(s') - u(s)\|}{\ln |s' - s|} &= \frac{\ln \|u(h_{w'}(S')) - u(h_w(S))\|}{n \ln(\frac{1}{2}) + O(1)} \\ &= \frac{\ln \|M_{w'}(u(S')) - M_w(u(S))\|}{n \ln(\frac{1}{2}) + O(1)}. \end{aligned} \quad (6)$$

Now two cases are possible:

- Suppose $w' = w$. Then by the previous Corollary, we have

$$\begin{aligned} \ln \|M_{w'}(u(S')) - M_w(u(S))\| &= \ln \|\vec{M}_w(u(S') - u(S))\| \\ &= \ln \|\vec{M}_w\| + \ln \|u(S') - u(S)\| + O(1). \end{aligned} \quad (7a)$$

- In the other case, by definition of n , w and w' are necessarily binary representations of two consecutive numbers. Without loss of generality, we may suppose that $s < s'$. Then we have $h_{w'}(0) = h_w(1)$, hence

$$\ln \|M_{w'}(u(S')) - M_w(u(S))\| = \ln \|\vec{M}_{w'}(u(S') - u(0)) + \vec{M}_w(u(1) - u(S))\|.$$

Now it is easy to see that, given two vectors \vec{u}, \vec{v} that can take values in a salient cone, the ratio between $\|\vec{u} + \vec{v}\|$ and $\|\vec{u}\| + \|\vec{v}\|$ is positively bounded. Since both terms in the above sum lie in \vec{K} , we can write, using once again the previous Corollary:

$$\begin{aligned} \ln \|M_{w'}(u(S')) - M_w(u(S))\| &= \\ &= \ln \left(\|\vec{M}_{w'}(u(S') - u(0))\| + \|\vec{M}_w(u(1) - u(S))\| \right) + O(1) \\ &= \max \left(\ln \|\vec{M}_{w'}(u(S') - u(0))\|, \ln \|\vec{M}_w(u(1) - u(S))\| \right) + O(1) \\ &= \max \left(\ln \|\vec{M}_{w'}\| + \ln \|u(S') - u(0)\|, \ln \|\vec{M}_w\| + \ln \|u(1) - u(S)\| \right) + O(1). \end{aligned}$$

Now note that we have

$$\vec{M}_w(\vec{v}_0) = \vec{M}_{w'}(\vec{v}_1).$$

Indeed, we may write $w = a_0 \dots a_k 0 1 \dots 1$ and $w' = a_0 \dots a_k 1 0 \dots 0$ (where both last groups of bits have equal, either zero or nonzero, length); the identity then follows from the equalities $\vec{M}_0(\vec{v}_0) = \frac{3}{5}\vec{v}_0$, $\vec{M}_1(\vec{v}_1) = \frac{3}{5}\vec{v}_1$ and $\vec{M}_0(\vec{v}_1) = \vec{M}_1(\vec{v}_0)$, that are straightforward to check (see also Figure 6). Since $\vec{v}_{0,1} \in \vec{K}$, we get

$$\ln \|\vec{M}_w\| + \ln \|\vec{v}_0\| + O(1) = \ln \|\vec{M}_{w'}\| + \ln \|\vec{v}_1\| + O(1),$$

or in other words, since $\|\vec{v}_0\| = \|\vec{v}_1\|$,

$$\ln \|\vec{M}_{w'}\| = \ln \|\vec{M}_w\| + O(1).$$

This allows us to write

$$\begin{aligned} \ln \|M_{w'}(u(S')) - M_w(u(S))\| &= \\ &= \ln \|\vec{M}_w\| + \max \left(\ln \|u(S') - u(0)\|, \ln \|u(1) - u(S)\| \right) + O(1). \end{aligned} \quad (7b)$$

Let us now check that in both cases, the middle term is bounded. Since $[0, 1]$ is compact and u is continuous, it is clearly bounded from above. For the other inequality, note that, in the first case, we have $(S' - S) = 2^n(s' - s) \geq \frac{1}{2} > \frac{1}{4}$; and in the second case, we have $(S' - 0) + (1 - S) = 2^n(s' - s) \geq \frac{1}{2}$, hence either $(S' - 0) \geq \frac{1}{4}$ or $(1 - S) \geq \frac{1}{4}$. Now consider the set

$$X := \left\{ (s_1, s_2) \in [0, 1]^2 \mid |s_1 - s_2| \geq \frac{1}{4} \right\}$$

and the function

$$\begin{aligned} X &\rightarrow \mathbb{R}_{\geq 0}, \\ (s_1, s_2) &\mapsto \|u(s_1) - u(s_2)\|. \end{aligned}$$

Since it is continuous and X is compact, it reaches its minimum value. But by Corollary 3, it cannot vanish; hence it is positively bounded. Thus $\|u(S') - u(S)\|$ in the first case, and $\max(\|u(S') - u(0)\|, \|u(1) - u(S)\|)$ in the second case, are positively bounded.

This allows us to merge (7a) with (7b), and to write:

$$\ln \|M_{w'}(u(S')) - M_w(u(S))\| = \ln \|\vec{M}_w\| + O(1),$$

and if we plug this into (6), the result follows. \square

5.2 Case of rational points

Theorem 12. *Let $s \in [0, 1] \cap \mathbb{Q}$. Then we know that the binary expansion of $s = 0.a_1a_2\dots$ is eventually periodic; let n be the length of its period, and $p = a_k\dots a_{k+n-1}$ its period. Then \vec{M}_p has distinct, positive, real eigenvalues, and we have*

$$\alpha(u, s) = \frac{\ln \lambda_p}{n \ln \frac{1}{2}},$$

where λ_p is the bigger eigenvalue of \vec{M}_p .

In practice, since we know the determinant of \vec{M}_p , we can calculate λ_p by the formula

$$\lambda_p = \frac{1}{2} \left(\text{tr } \vec{M}_p + \sqrt{(\text{tr } \vec{M}_p)^2 - \left(\frac{3}{25}\right)^n} \right).$$

Note that Proposition 7 also allows us to infer $u'(s)$. Indeed, we shall see in the next section that we never have $\alpha(u, s) = 1$.

To prove the theorem, we shall need the following lemma:

Lemma 13. *Let $s \in [0, 1]$, let $w = a_1 \dots a_n$ be a word on the alphabet $\{0, 1\}$. Then we have:*

$$\alpha_{\inf}(u, h_w(s)) = \alpha_{\inf}(u, s);$$

$$\alpha_{\sup}(u, h_w(s)) = \alpha_{\sup}(u, s).$$

Proof. Let $0.b_1b_2\dots$ be the binary expansion of s ; let $\vec{u}_0 \in \vec{K}$. It is well-known that $\frac{\|\vec{M}_w(\vec{u})\|}{\|\vec{u}\|}$ is bounded when \vec{u} varies; since \vec{M}_w is invertible, the reciprocal of that quantity is bounded as well; hence the quantity

$$\ln \|\vec{M}_{a_1\dots a_kb_1\dots b_n}(\vec{u}_0)\| - \ln \|\vec{M}_{b_1\dots b_n}(\vec{u}_0)\|$$

is bounded. The result follows immediately from Proposition 9. \square

Proof of Theorem 12. The lemma we have just shown allows us to drop the preperiod of the expansion of s : we may suppose that it is periodic. From the formula (5) it follows that we have:

$$\alpha_{\inf}(u, s) = \alpha_{\sup}(u, s) = \lim_{k \rightarrow \infty} \frac{\ln \|\vec{M}_p^k\|}{kn \ln \left(\frac{1}{2}\right)}$$

Now since \vec{K} is stable by \vec{M}_p , we can also say that \tilde{K} is stable by \tilde{M}_p . Since \tilde{M}_p is contractive, this means that it has exactly one fixed point in \tilde{K} , which means that \vec{M}_p has a unique eigenvector \vec{v}_p lying in \vec{K} and associated with a real positive eigenvalue λ_p (we will justify in a minute that it is indeed the bigger one). Let μ_p be the other eigenvalue; since $\det(\vec{M}_p) = (\frac{3}{25})^n > 0$, it is also real positive, and thus the associated eigenvector \vec{w}_p cannot lie in \vec{K} . Using once again the fact that $\vec{M}_p(\vec{K})$ is a proper subset of \vec{K} and considering the positions of the relevant vectors relative to each other and to \vec{K} , it can easily be seen that $\lambda_p > \mu_p$. It follows that $\|\vec{M}_p^k\| \sim \lambda_p^k$, hence the result. \square

Example 1. It follows from Theorem 12 that whenever s is dyadic rational, we have $\alpha(u, s) = \frac{\ln \frac{3}{5}}{\ln \frac{1}{2}} \approx 0,737$ (indeed, in this case, λ_p is simply the bigger eigenvalue of \vec{M}_0 or \vec{M}_1 , namely $\frac{3}{5}$). In particular, we have in this case $u'(s) = \infty$.

Note that Lemma 13 is interesting in itself, even outside the case $s \in \mathbb{Q}$: it shows that changing a finite number of bits in the binary expansion of s does not change its Hölder exponent. For completeness' sake, let us also mention the following (fairly obvious) symmetry of u . Let

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(in Figure 5, it simply corresponds to the vertical symmetry axis). Then

Proposition 14. *For all $s \in [0, 1]$, we have $u(1-s) = Pu(s)$.*

Corollary 15. *For all $s \in [0, 1]$, $\alpha_{\inf}(u, 1-s) = \alpha_{\inf}(u, s)$ and $\alpha_{\sup}(u, 1-s) = \alpha_{\sup}(u, s)$.*

Proof. Clearly, $P^2 = 1$, $M_1 = PM_0P^{-1} = PM_0P$, $M_0 = PM_1P$, and K is stable by P . Let $s = 0.a_1a_2\dots$ be its binary expansion; then, simply using (3), we have:

$$\begin{aligned} u(1-s) &= u(0.\bar{a}_1\bar{a}_2\dots) \\ &= \lim_{n \rightarrow \infty} M_{\bar{a}_1\dots\bar{a}_n}u_0 \\ &= \lim_{n \rightarrow \infty} PM_{a_1\dots a_n}Pu_0 \\ &= Pu(s), \end{aligned}$$

where \bar{a} is shorthand for $1-a$ and $u_0 \in K$ is an arbitrary starting point. \square

5.3 Derivative of u

Theorem 16. *For all $s \in [0, 1]$, $u'(s)$ is either 0, ∞ or not defined at all.*

Proof. Assume the opposite: let s be a number in $[0, 1]$ such that $u'(s)$ exists, is finite and positive. Then we have, in particular:

$$\lim_{t \rightarrow s} \frac{\|u(t) - u(s)\|}{|t - s|} = C$$

for some finite positive constant C .

Let $s = 0.a_1a_2\dots$ be a binary expansion. Let t_n be a sequence, taking values in $[0, 1]$, such that for all $n \in \mathbb{N}$, the binary expansion t_n coincides with that of s in the first n places. Obviously, such a sequence converges to s . We then have, for all $n \in \mathbb{N}$:

$$\frac{\|u(t_n) - u(s)\|}{|t_n - s|} = \frac{\|\vec{M}_{a_1\dots a_n}(u(T_n) - u(S_n))\|}{2^{-n}|T_n - S_n|},$$

where $S_n = h_{a_1\dots a_n}^{-1}(s) = 0.a_{n+1}a_{n+2}\dots$ and $T_n = h_{a_1\dots a_n}^{-1}(t_n)$. Since $[0, 1]$ is compact, we may extract from the natural numbers a subsequence $\phi(n)$ such that $S_{\phi(n)}$ converges to some value S_∞ .

Suppose that $S_\infty < 1$ (by reversing the order of s and t , we can similarly treat the case $S_\infty > 0$). From now on, we shall require that $t_n > s$. By Corollary 11, we know that

$$\frac{\|\vec{M}_{a_1\dots a_n}\| \|u(T_n) - u(S_n)\|}{\|\vec{M}_{a_1\dots a_n}(u(T_n) - u(S_n))\|}$$

is positively bounded. We also know that

$$\frac{\|\vec{M}_{a_1\dots a_n}(u(T_n) - u(S_n))\|}{2^{-n}|T_n - S_n|}$$

tends to C , hence it is positively bounded as well. Suppose for a moment that t_n is such that T_n is identically equal to 1. Then for sufficiently large n , $S_{\phi(n)}$ varies in a closed interval that does not contain 1, and since u is continuous and injective, $\frac{\|(u(T_{\phi(n)}) - u(S_{\phi(n)}))\|}{|T_{\phi(n)} - S_{\phi(n)}|}$ is positively bounded.

This means that $\frac{\vec{M}_{a_1 \dots a_{\phi(n)}}}{2^{-\phi(n)}}$ is positively bounded in norm, so that it takes values in a compact space. Hence from $\phi(n)$ we can extract yet another subsequence, that we shall call $\psi(n)$, such that $\frac{\vec{M}_{a_1 \dots a_{\psi(n)}}}{2^{-\psi(n)}}$ converges; let us call \vec{M}_∞ its limit. Note that, though we used a particular value of t_n to prove that this construction was possible, the actual objects that we have constructed do not depend on it.

Now let T be any point in $(S_\infty, 1]$; we set $t_n = h_{a_1 \dots a_n}(T)$, so that for all n , $T_n = T$. We then have:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\|u(t_{\psi(n)}) - u(s)\|}{|t_{\psi(n)} - s|} &= \lim_{n \rightarrow \infty} \frac{\|\vec{M}_{a_1 \dots a_{\psi(n)}}(u(T) - u(S_{\psi(n)}))\|}{2^{-\psi(n)}|T - S_{\psi(n)}|} \\ &= \frac{\|\vec{M}_\infty(u(T) - u(S_\infty))\|}{|T - S_\infty|}, \end{aligned}$$

which means that the latter ratio is equal to C regardless of the value of T .

This claim is definitely too strong to be true, and we shall soon prove it is absurd. Note that we have

$$\det \left(\frac{\vec{M}_{a_1 \dots a_{\psi(n)}}}{2^{-\psi(n)}} \right) = \det \left(2\vec{M}_i \right)^{\psi(n)} = \left(\frac{12}{25} \right)^{\psi(n)},$$

which tends to 0 as n increases. Hence $\det(\vec{M}_\infty) = 0$, in other words, \vec{M}_∞ has rank 1. The claim can then be reformulated as follows: there exists some linear form Φ such that the ratio

$$\frac{\Phi(u(T) - u(S_\infty))}{T - S_\infty}$$

is constant. The interval (S_∞, T) contains infinitely many dyadic rational points; by Corollary 6, we may find in this interval a dyadic rational value T' such that $\Phi(\vec{D}u(T')) \neq 0$. But Example 1 together with Lemma 8 tell us that in this case, $\Phi(u(T) - u(S_\infty))$ as a function of T has an infinite derivative in T' . So must the whole ratio: contradiction.

Now let $s \in [0, 1] \cap \mathbb{Q}$, and suppose $u'(s)$ is undefined. Let us apply Theorem 12: it is only possible if $\alpha(u, s) = 1$. Let p be the period of the binary expansion of s , n the length of the period; then we have $\lambda_p = (\frac{1}{2})^n$. We know that M_p has determinant $(\frac{3}{25})^n$ and that one of its eigenvalues is equal to 1. Hence $\text{tr } M_p = 1 + (\frac{1}{2})^n + (\frac{6}{25})^n$, and $\text{tr } 25^n M_p = 25^n + 6^n + (\frac{25}{2})^n$. On the other hand, we know that in the canonical basis, $25M_0$ and $25M_1$ are given by matrices with integer coefficients, which means that so is $25^n M_p$, hence $\text{tr } 25^n M_p$ must be integer. Contradiction. \square

Corollary 17. *For Lebesgue-almost all $s \in [0, 1]$, $u'(s) = 0$.*

Proof. Let Φ be a linear form whose kernel does not meet \vec{K} ; then from Corollary 6, it is easy to deduce that the function $\Phi \circ u$ is monotonic. But it is a well-known result (see for example [3], Theorem 8.19) that a monotonic function is differentiable (in its usual sense of “has a finite derivative”) almost everywhere. Since by Lemma 8, $\Phi \circ u$ behaves in the same way as u itself, the result follows. \square

5.4 Numerical values and estimations of Hölder exponent

Here is a table of all possible periods shorter than or equal to 7 (up to cyclic permutation), ordered by Hölder exponent:

s	Period p	Length n	$5^n \operatorname{tr} \vec{M}_p$	Approximate value of α
$\frac{1}{3}$	01	2	7	1.119
$\frac{21}{127}$	0010101	7	388	1.096
$\frac{11}{63}$	001011	6	175	1.086
$\frac{5}{31}$	00101	5	76	1.085
$\frac{1}{5}$	0011	4	34	1.078
$\frac{19}{127}$	0010011	7	436	1.072
$\frac{11}{127}$	0001011	7	472	1.055
$\frac{13}{127}$	0001101	7	472	1.055
$\frac{1}{7}$	001	3	16	1.050
$\frac{3}{31}$	00011	5	88	1.040
$\frac{5}{63}$	000101	6	211	1.039
$\frac{1}{9}$	000111	6	223	1.025
$\frac{9}{127}$	0001001	7	580	1.012
$\frac{5}{127}$	0000101	7	616	0.999
$\frac{1}{21}$	000011	6	250	0.997
$\frac{7}{127}$	0000111	7	628	0.995
$\frac{1}{15}$	0001	4	43	0.982
$\frac{3}{127}$	0000011	7	736	0.962
$\frac{1}{31}$	00001	5	124	0.936
$\frac{1}{63}$	000001	6	367	0.903
$\frac{1}{127}$	0000001	7	1096	0.880
0	0	1	4	0.737

When we look at this table, it becomes apparent that globally, numbers whose binary expansions contain long clusters of equal consecutive bits tend to have smaller Hölder exponents, while those whose binary expansions are more “mixed” tend to have bigger Hölder exponents. The exponent is lowest when the binary expansion ends in an unbroken string of zeroes or ones (that is, for dyadic rational numbers — see below), and seems highest when zeroes and ones strictly alternate (that is, for $\frac{1}{3}$ and its dyadic-rational multiples; the author does not know how to prove this). The author was able to prove the following statement, that partly confirms this intuition:

Proposition 18. *Let $s \in [0, 1]$, $s = 0.a_1a_2\dots$ its binary expansion. Let us define*

$$d_{\inf, \sup}(s) := \liminf_{n \rightarrow \infty} \sup \frac{1}{n} \sum_{i=1}^n |a_i - a_{i+1}|.$$

We then have:

$$\alpha_{\inf}(u, s) \leq \frac{\ln \frac{3}{5}}{\ln \frac{1}{2}} + d_{\inf}(s);$$

$$\alpha_{\sup}(u, s) \leq \frac{\ln \frac{3}{5}}{\ln \frac{1}{2}} + d_{\sup}(s).$$

Informally, $d_{\inf}(s)$ and $d_{\sup}(s)$ give bounds for the asymptotic proportion of places where the bit changes value; the latter can also be understood as the reciprocal of the asymptotic average length of blocs of consecutive equal bits.

Corollary 19. *If $d_{\sup}(s) < 1 - \frac{\ln \frac{3}{5}}{\ln \frac{1}{2}} \approx 0.263 \approx \frac{1}{3.802}$, then $u'(s) = \infty$.*

To prove the proposition, we will need a small technical lemma:

Lemma 20. *For any $\vec{u} \in \vec{K}$, $n \in \mathbb{N}$ and $i \in \{0, 1\}$, we have*

$$\frac{\|\vec{M}_i^n(\vec{u})\|_2}{\|\vec{u}\|_2} \geq \frac{1}{2} \left(\frac{3}{5}\right)^n.$$

Proof. Let us first express \vec{K} in the bases that diagonalise \vec{M}_i . A quick calculation shows that we have:

$$\vec{K} = \left\{ a\vec{v}_0 + b\vec{w}_0 \mid 0 \leq \frac{b}{a} \leq 3 \right\} = \left\{ a\vec{v}_1 + b\vec{w}_1 \mid 0 \geq \frac{b}{a} \geq -3 \right\}$$

Let $i \in \{0, 1\}$ and $\vec{u} = a\vec{v}_i + b\vec{w}_i \in \vec{K}$. We then have $\vec{M}_i(\vec{u}) = \frac{3}{5}a\vec{v}_i + \frac{1}{5}b\vec{w}_i$. Now let $n \in \mathbb{N}$; we have (since \vec{v}_i and \vec{w}_i are orthogonal):

$$\begin{aligned} \left(\frac{\|\vec{M}_i^n(\vec{u})\|_2}{\|\vec{u}\|_2} \right)^2 &= \frac{\left\| \left(\frac{3}{5}\right)^n a\vec{v}_i + \left(\frac{1}{5}\right)^n b\vec{w}_i \right\|_2^2}{\|a\vec{v}_i + b\vec{w}_i\|_2^2} \\ &= \frac{\left(\frac{3}{5}\right)^{2n} \|a\vec{v}_i\|_2^2 + \left(\frac{1}{5}\right)^{2n} \|b\vec{w}_i\|_2^2}{\|a\vec{v}_i\|_2^2 + \|b\vec{w}_i\|_2^2} \\ &\geq \left(\frac{3}{5}\right)^{2n} \frac{\|a\vec{v}_i\|_2^2}{\|a\vec{v}_i\|_2^2 + \|b\vec{w}_i\|_2^2} \\ &= \left(\frac{3}{5}\right)^{2n} \frac{1}{1 + \left(\frac{b}{a}\right)^2 \frac{\|\vec{w}_i\|_2^2}{\|\vec{v}_i\|_2^2}} \\ &= \left(\frac{3}{5}\right)^{2n} \frac{1}{1 + \frac{1}{3} \left(\frac{b}{a}\right)^2} \\ &\geq \left(\frac{3}{5}\right)^{2n} \frac{1}{1 + \frac{1}{3} 3^2} \\ &= \left(\frac{1}{2} \left(\frac{3}{5}\right)^n\right)^2. \end{aligned}$$

□

Proof of Proposition 18. Let $\vec{u}_0 \in \vec{K}$; by Proposition 9 we may then write

$$\alpha_{\inf}(u, s) = \liminf_{n \rightarrow \infty} \frac{\ln \|\vec{M}_{a_1 \dots a_n}\|}{n \ln \left(\frac{1}{2}\right)} = \liminf_{n \rightarrow \infty} \frac{\ln \|\vec{M}_{a_1 \dots a_n}(\vec{u}_0)\|_2}{n \ln \left(\frac{1}{2}\right)}.$$

We may now decompose the binary expansion of s into clusters of consecutive equal bits. Let $n \in \mathbb{N}$; we have:

$$\vec{M}_{a_1 \dots a_n} = \vec{M}_{a_1} \dots \vec{M}_{a_n} = \vec{M}_{\alpha_1}^{l_1} \vec{M}_{\alpha_2}^{l_2} \dots \vec{M}_{\alpha_{k(n)-1}}^{l_{k(n)-1}} \vec{M}_{\alpha_{k(n)}}^{r(n)},$$

where the values of α_i alternate between 0 and 1 (so that $\alpha_1 = a_1$, $\alpha_2 = 1 - a_1$, $\alpha_3 = a_1$, \dots , $\alpha_{k(n)} = a_n$) and all l_i are positive. From Lemma 20, it follows that

$$\|\vec{M}_{a_1 \dots a_n}(\vec{u}_0)\|_2 \geq \left(\frac{1}{2}\right)^{k(n)} \left(\frac{3}{5}\right)^n \|\vec{u}_0\|_2$$

hence (don't forget that $\ln\left(\frac{1}{2}\right) < 0$):

$$\frac{\ln \|\vec{M}_{a_1 \dots a_n}(\vec{u}_0)\|_2}{n \ln\left(\frac{1}{2}\right)} \leq \frac{k(n)}{n} + \frac{\ln\left(\frac{3}{5}\right)}{\ln\left(\frac{1}{2}\right)} + \frac{\ln \|\vec{u}_0\|_2}{n}.$$

When n grows, the third term tends to 0, and the first term's lower limit is by definition equal to $d_{\inf}(x)$, hence the conclusion. \square

Finding lower bounds for the Hölder exponent seems much harder than finding upper bounds. Here is a trivial one: we obviously have

$$\forall \vec{u} \in \vec{\mathcal{H}}, \forall i \in \{0, 1\}, \quad \|M_i(\vec{u})\| \leq \frac{3}{5} \|\vec{u}\|,$$

hence the Hölder exponent is always greater or equal to $\frac{\ln\left(\frac{3}{5}\right)}{\ln\left(\frac{1}{2}\right)} \approx 0.737$. However, this is useless for finding points where the derivative is zero. The author suspects that if the binary expansion of s never contains more than two equal consecutive bits, then $u'(s) = 0$ — but it has yet to be checked.

References

- [1] A. A. Kirillov, *A tale of two fractals*, <http://www.math.upenn.edu/~kirillov/MATH480-F07/tf.pdf>
- [2] R. S. Strichartz, *Differential equations on fractals: A tutorial*, Princeton University Press, 2006
- [3] W. Rudin, *Real and complex analysis*, Mc Graw-Hill, 1966.